CENG 222

Statistical Methods for Computer Engineering

Spring '2016-2017

Assignment 4

Deadline: May 26, 23:59
Submission: via COW

Student Information

Full Name : Murat TOPAK
Id Number : 2036218

Answer 9.8

a) Use the formula 9.5 given in the textbook. Confidence interval for the mean when σ known is $\bar{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$. All the parameters we need are given in the question text. Also, $z_{\alpha/2} = z_{0.025} = 1.96$ from Table A4. Thus, the interval is $42 \pm 1.96(\frac{5}{\sqrt{64}}) = [40.775, 43.225]$.

b) Now we are also given population mean installation time $\mu = 40$ minutes. Let X be the random variable denoting the installation time on my PC.
We want to find $P(40.775 \leq X \leq 43.225)$. Assuming X is a normal random variable, we standardize the inequality and turn it into to find

$$P(\frac{40.775 - \mu}{\sigma} \leq X \leq \frac{43.225 - \mu}{\sigma}) = P(0.645 \leq Z \leq 0.155) = 0.7406 - 0.5616 = 0.1790$$

In the last step a standard normal distribution table has been used to find values 0.7406 and 0.5616

Answer 9.16

a) Use the formula given on the page 257 to calculate the confidence interval for the difference of proportions. It is $\hat{p_1} - \hat{p_2} \pm z_{\alpha/2}\sqrt{\frac{\hat{p_1}(1-\hat{p_1})}{n_1} + \frac{\hat{p_2}(1-\hat{p_2})}{n_2}}$ where

$\hat{p_1} = 10/250 = 0.04$, $\hat{p_2} = 18/300 = 0.06$, $n_1 = 250$, $n_2 = 300$. Also note that $z_{\alpha/2} = z_{0.01} = 2.33$ from Table A4. When these values are substituted, we have $(-0.063, 0.023)$ for the confidence interval.

b) We should test the null hypothesis $H_0 : p_1 = p_2$ against its two sided alternative $H_a : p_1 \neq p_2$ at the $\alpha = 2\%$ level. According to the duality between two sided tests and confidence intervals, we cannot reject the null hypothesis as 98% confidence interval for $p_1 - p_2$ contains zero. Thus, there is no significant difference between two lots.

Answer 10.3

k	1	2	3	4	5
B_k	$(-\infty, -1.073]$	$(-1.073, -0.386]$	$(-0.386, 0.218]$	$(0.218, 0.969]$	$(0.969, \infty)$
p_k	0.1772	0.2051	0.217	0.2262	0.1745
$Exp(k)$	17.72	20.51	21.70	22.62	17.45
$Obs(k)$	20	20	20	20	20

To apply the goodness of fit test, I have set up the boundaries of intervals so there are 5 bins. Method of maximum likelihood estimation for normal distribution returs $\hat{\mu} = \bar{X}$, and $\hat{\sigma} = \sqrt{\frac{(n-1)s^2}{n}}$.

The sample in the question consists of 100 entries. With the help of an online tool, the sample mean $\bar{X}$ is equal to -0.05773 and the sample variance s^2 is equal to 1.119. Then the normal distribution's estimated parameters are $\hat{\mu} = -0.05773$, and $\hat{\sigma} = 1.0963$. You can find the area under the normal curve(p_k) for each interval with respect to these estimated parameters. $Exp(k) = np_k$, and we have observed 20 items for each interval in our sample.

Chi square statistic is now calculated according to the formula 10.1. It is 1.115.